



Analysis of Spatiotemporal Variations and Driving Mechanisms of Ecological Environment Quality in Yantai City Based on the XGBoost-SHAP Model

Jinli Zhou, Ruei-Yuan Wang*

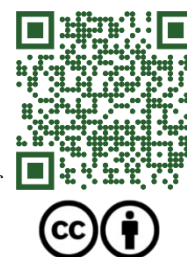
School of Sciences, Guangdong University of Petrochem Technology (GDUPT), Maoming 525000, China

Received: 30 Mar 2026; Received in revised form: 27 Apr 2026; Accepted: 03 May 2026; Available online: 13 May 2026

©2026 The Author(s). Published by Infogain Publication. This is an open-access article under the CC BY license

(<https://creativecommons.org/licenses/by/4.0/>).

Abstract— Conducting long-term spatiotemporal monitoring of ecological environment quality and identifying its driving mechanisms are of great significance for coordinating ecological protection with high-quality regional economic development. Based on Landsat remote sensing data from 2005 to 2025 on the Google Earth Engine (GEE) platform, this study employed Principal Component Analysis (PCA) to construct the Remote Sensing Ecological Index (RSEI), analyzing the spatiotemporal evolution patterns of RSEI in Yantai City. The study further utilized the explainable extreme gradient boosting model (XGBoost-SHAP) to reveal key natural and anthropogenic drivers of ecological environment quality, along with their nonlinear and interactive effects. The results indicate that: (1) On a temporal scale, the RSEI of Yantai City exhibited a fluctuating trend of initial decline followed by an increase, then another decline and subsequent rise, with a multi-year average of 0.438, indicating an overall moderate level of ecological environment quality; (2) Spatially, the RSEI in the study area showed significant heterogeneity, with high-value areas concentrated in central Yantai and low-value areas distributed near the coast; (3) Ecological quality was primarily driven by land use type and precipitation, with anthropogenic factors also playing a significant role. This research provides methodological references and scientific foundations for ecological monitoring, risk assessment, and territorial spatial optimization in coastal cities.



Keywords— Remote Sensing Ecological Index (RSEI), XGBoost-SHAP Model, Google Earth Engine (GEE), Principal Component Analysis (PCA), Ecological Environment Quality (EEQ), Yantai City

I. INTRODUCTION

Environmental degradation caused by human activities, along with intensive urbanization, has posed severe crises and challenges to global ecosystems (Yang et al., 2024). The ecological environment serves as a critical foundation for the sustainable development of human society, directly influencing regional economic growth potential and residents' well-being (Liu et al., 2024). In the

face of these pressing ecological challenges, effectively monitoring environmental changes has become a key priority in current conservation efforts. Remote sensing technology, with its advantages of large-scale, rapid, comprehensive, and continuous observation, enables holistic monitoring of ecological changes (Geng et al., 2022). Building on this, Xu Hanqiu (2013) proposed the Remote Sensing Ecological Index (RSEI) (Xu, 2013),

which integrates multiple indicator factors based on remote sensing data. It identifies greenness, wetness, heat, and dryness as four crucial indicators linked to human survival, development, and subjective perception of ecological conditions. By synthesizing various ecological factors, RSEI provides a quantitative method for assessing ecosystem health and environmental quality, now widely applied for rapid and objective evaluation of regional ecological conditions (Huang et al., 2024).

Moreover, long-term ecological environment quality (EEQ) monitoring is essential for a comprehensive understanding of environmental changes and for providing substantial data to support decision-making. However, this approach heavily relies on large-scale, high-frequency remote sensing data, and traditional assessment methods are often constrained by bottlenecks in acquiring and preprocessing massive remote sensing datasets. Google Earth Engine (GEE), with its integrated multi-source data resources and powerful cloud computing capabilities, effectively addresses the time-consuming and labor-intensive challenges of conventional methods. The platform eliminates the need for local downloads and complex preprocessing steps such as radiometric calibration and atmospheric correction, enabling efficient large-scale time-series analysis. It has been widely applied in regional long-term ecological environment quality research (Wang et al., 2025; Zhang et al., 2025).

Accurately identifying the key drivers of ecological environment quality changes is fundamental to scientifically formulating conservation and restoration strategies and implementing targeted measures. Current studies commonly employ methods such as multiple regression analysis (Wei et al., 2021), structural equation modeling (Xu et al., 2025), geographically weighted regression (Sun et al., 2024), and geographical detectors (Zhang et al., 2024). In contrast, machine learning approaches, with their strong nonlinear fitting and high-dimensional data processing capabilities, offer a new research pathway for identifying and attributing RSEI driving factors. Among these, the extreme gradient boosting (XGBoost) algorithm combined with Shapley additive explanations (SHAP) models establishes an

analytical paradigm that balances predictive performance and interpretability, significantly outperforming traditional methods (Bao et al., 2025). This approach has been widely adopted across numerous fields.

Currently, there is limited research on ecological environment monitoring in coastal cities. Therefore, this study takes the coastal city of Yantai as an example, using the RSEI index to conduct ecological monitoring and analysis of the area. It constructs and calculates the remote sensing ecological index of Yantai City to promptly and objectively assess the quality of the ecological environment and its spatiotemporal distribution characteristics. The aim is to provide a scientific basis for urban construction and ecological environmental protection under the context of rapid urbanization (Yuan et al., 2022).

II. STUDY AREA AND DATA SOURCES

2.1 Study Area

Yantai City is located in the eastern part of the Shandong Peninsula (between E121°57'~E119°34' and N38°23'~N36°16'), situated between the two major economic zones of Beijing-Tianjin-Hebei and the Yangtze River Delta, serving as a crucial port city in the Bohai Rim region (Figure 1). The city administers five municipal districts—Zhifu, Fushan, Muping, Laishan, and Penglai—as well as six county-level cities: Longkou, Laiyang, Laizhou, Zhaoyuan, Qixia, and Haiyang, boasting significant geographical advantages. The region is predominantly characterized by low mountains and hills, featuring a composite landform of mountains, hills, and plains, with distinct spatial variations in hydrothermal conditions and underlying surface types.

As a typical area of land-sea interaction and rapid urbanization, its ecosystem is driven by both natural evolution and human activities, resulting in prominent spatiotemporal heterogeneity in environmental quality. This study employs the RSEI and XGBoost models to dynamically assess ecological environment quality, providing an interpretative method for identifying the evolution of ecological quality and its driving mechanisms in the region.

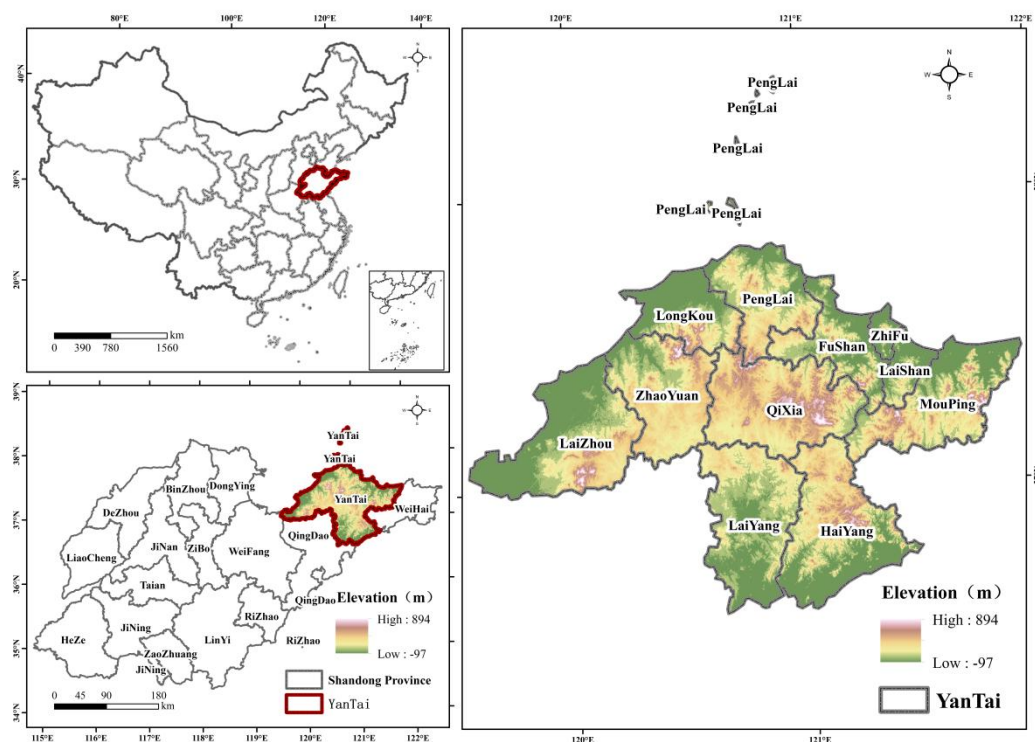


Fig.1 Topographic Overview Map of Yantai City

2.2 Data Sources

The data used in this study can be broadly categorized into two types: (1) RSEI data generated from remote sensing imagery and (2) various data sources for driving factor analysis, as detailed below.

(1) RSEI Data Generated from Remote Sensing Imagery

The Landsat remote sensing imagery data from the Google Earth Engine (GEE) platform consists of Tier 1 (T1) Surface Reflectance (SR) products (Wang et al., 2023). The remote sensing data involved include surface reflectance data from Landsat 5 (TM) and Landsat 8 (OLI), with a spatial resolution of 30 meters. These datasets have undergone preprocessing steps such as radiometric correction, atmospheric correction, and geometric correction.

GEE provides a robust cloud-based platform capable of processing and analyzing large-scale remote sensing datasets, making RSEI computation faster and more convenient. Additionally, GEE's parallel cloud computing capabilities significantly enhance image processing efficiency, addressing issues of data gaps caused by cloud cover. This platform also holds potential for expanding the

application of RSEI in large-scale, long-term ecological monitoring.

Based on the aforementioned Landsat SR data, the four core component indicators for calculating RSEI were extracted and computed in GEE: Normalized Difference Vegetation Index (NDVI), Normalized Difference Bare Soil and Built-up Index (NDBSI), Wetness Component (WET), and Land Surface Temperature (LST). Subsequently, principal component analysis (PCA) was applied to integrate these four indicators, constructing the comprehensive EEQ index RSEI. To ensure the scientific validity and reliability of the inversion results, the selection and correction of image data adhered to strict standards. Considering the impact of vegetation phenological differences on image data, this study exclusively utilized images from the peak vegetation growth season (June to October) (Li et al., 2025), with water bodies masked (threshold value of 0.05).

(2) Various Data Sources for Driving Factor Analysis

The formation of regional ecological environments is the result of interactions between natural factors and human activities. Drawing on previous research and considering the actual conditions of the study area, ten

indicators were selected from two dimensions—natural and human activities. These include:

X1: Slope (SLO), X2: Elevation (DEM), X3: Population distribution (POP), X4: Annual precipitation (PRE), X5: GDP, X6: Annual mean air temperature (AT), X7: Nighttime light intensity (NTL), Under the Land use type (China Land Cover Dataset, CLCD): X8: Built-up land (BL), X9: Grassland (GL), X10: Forest land (FL).

These indicators are used to explore the influencing mechanisms of RSEI in the study area for each year, as detailed in Table 1. Among them,

- A. NDVI, NDBSI, WET, and LST were derived from calculations on the GEE platform.
- B. AT data was sourced from the China Meteorological Data Network, with a spatial resolution of 1000m; used to explore the driving factors of RSEI.
- C. PRE data was obtained from the National Tibetan Plateau Data Center, with a spatial resolution of 1000m; used to explore the driving factors of RSEI.
- D. DEM data was sourced from the Geospatial Data Cloud, with a spatial resolution of 30m; used to explore the driving factors of RSEI.

- E. Slope data was calculated from DEM, with a spatial resolution of 30m; used to explore the driving factors of RSEI.
- F. Land use data was derived from Wuhan University's China 30m Annual Land Cover Dataset (CLCD), with a spatial resolution of 30m; used to explore the driving factors of RSEI.
- G. NTL data was sourced from the National Oceanic and Atmospheric Administration (NOAA), with a spatial resolution of 1000m; used to explore the driving factors of RSEI.
- H. POP data was obtained from the Oak Ridge National Laboratory's LandScan Global Population Dataset, with a spatial resolution of 1000m; used to explore the driving factors of RSEI.

For data preprocessing, the fishnet tool in ArcGIS 10.8 was used to create a 1km × 1km grid, followed by processing natural factor data (including AT, PRE, DEM, SLO, and CLCD) and human activity factor data (including NTL, GDP, and POP) using the "Extract Multi Values to Points" tool for subsequent RSEI influencing factor analysis. All geographic data in this study were uniformly projected in the WGS_1984_UTM_Zone_50N coordinate system.

Name	Source	Purpose
NDVI	Generated from running on the GEE platform (https://developers.google.cn/earth-engine?hl=zh-cn)	For subsequent RSEI calculations
NDBSI		
WET		
LST		
Temperature(AT)	China Meteorological Data Network (http://data.cma.cn/)	To explore the driving factors of RSEI
Average annual precipitation(PRE)	Qinghai-Tibet Plateau Scientific Data Center (https://data.tpdac.ac.cn/)	
Elevation(DEM)	Geospatial Data Cloud (https://www.gscloud.cn/)	
Slope (SLO)	Calculated from DEM	
Land use(CLCD)	Wuhan University's CLCD China 30m Annual Land Cover Dataset (https://doi.org/10.5281/zenodo.18180184)	
Population(POP)	Oak Ridge National Laboratory, USA, LandScan Global Population Dataset (https://landscan.ornl.gov/)	

GDP	China GDP Spatial Distribution Dataset (https://github.com/thestarlab/ChinaGDP)
Nighttime light data (NTL)	Earth Observation Group under the National Oceanic and Atmospheric Administration (NOAA) of the United States (Earth Observation Group, EOG) (https://payneinstitute.mines.edu/eog/)

Table 1: Required Data and Sources

III. METHODOLOGY

3.1 Method

This study utilized the GEE platform to acquire remote sensing ecological indicator data, including NDVI, WET, NDBSI, and LST metrics. These indicators were synthesized into an ecological index (RSEI) through PCA to interpret the spatiotemporal distribution patterns after grading. The XGBoost model was employed to establish nonlinear response relationships between ecological quality and driving factors. Key driving mechanisms were then analyzed using SHAP importance analysis and PDP (Partial Dependence Plot). Finally, a coupled analysis of the results was conducted to derive scientific evidence for the study area (Figure 2).

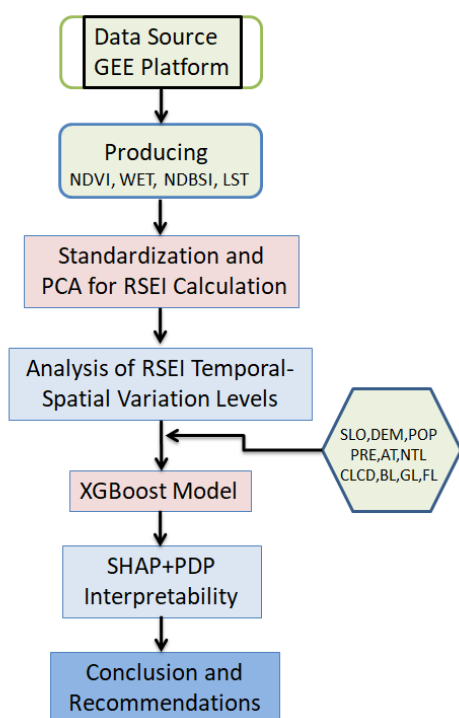


Fig.2 The Technical Roadmap of the Study

3.2 Construction Principle of RSEI

RSEI is a comprehensive model for quantitatively assessing regional ecological environment quality based on

multi-source remote sensing data. This model integrates four key ecological parameters: greenness (NDVI), wetness (WET), dryness (NDBSI), and heat (LST). It converts the values of these factors into dimensionless index values ranging from [0,1], where higher values indicate better regional ecological environment quality (Zhang et al., 2026).

Referencing the five-level classification concept from the "Technical Criterion for Eco-Environmental Assessment" (HJ/T192-2006) and following the RSEI grading convention established by Xu Hanqiu et al. (2013), the normalized RSEI values (0-1) are divided into five levels at intervals of 0.2: 0.8-1.0 (Excellent), 0.6-0.8 (Good), 0.4-0.6 (Moderate), 0.2-0.4 (Fair), and 0-0.2 (Poor) (Zhou et al., 2023).

The Normalized Difference Vegetation Index (NDVI) is one of the core indicators of RSEI, calculated as shown in Equation (1).

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (1)$$

Here, NIR represents the near-infrared band reflectance, and R is the red band reflectance. NDVI utilizes the strong absorption of the red band and the high reflectance of the near-infrared band by vegetation to measure vegetation coverage and growth vitality. A higher NDVI indicates denser vegetation, reflecting stronger ecosystem functions such as photosynthetic production and soil-water conservation. Conversely, a lower NDVI suggests sparser vegetation, which may indicate issues like bare land expansion, urban development, or ecological degradation.

The Wetness Index (WET) is constructed based on the Tasseled Cap Wetness (TCW) transformation, with tailored algorithms for different sensors (e.g., the Landsat series). Taking Landsat 8 as an example, WET is derived from a linear combination of reflectance values in the blue, green,

red, near-infrared, shortwave infrared 1, and shortwave infrared 2 bands. The formula is (2):

$$WET = 0.1511\rho_{Blue} + 0.1973\rho_{Green} + 0.3283\rho_{Red} + 0.3407\rho_{NIR} - 0 \quad (2)$$

The Wetness Index (WET) simulates the moisture content of an ecosystem. High values correspond to wet surface conditions (such as water bodies, wetlands, and areas with high vegetation coverage), indicating good water retention and regulation in wetland systems. Low values correspond to dry conditions (such as arid regions and built-up areas), representing water stress in ecosystems or disruption of the hydrological cycle due to human interference.

The Dryness Index employs the Normalized Difference Built-up and Barren Index (NDBSI) to focus on surface desiccation and human construction disturbances, calculated using formulas (3), (4), and (5).

$$NDBSI = \frac{(SI + BI)}{2} \quad (3)$$

$$SI = \frac{\rho_{SWIR1} + \rho_{RED} - \rho_{NIR} - \rho_{Blue}}{\rho_{SWIR1} + \rho_{RED} + \rho_{NIR} + \rho_{Blue}} \quad (4)$$

$$SI = \frac{\rho_{SWIR1} - \rho_{NIR}}{\rho_{SWIR1} + \rho_{NIR}} \quad (5)$$

The NDBSI integrates bare soil and construction information, with high values corresponding to built-up areas and bare land, reflecting the compression of ecological space and surface desiccation caused by human development, which are typical indicators of ecological degradation. Low values are associated with ecological land such as vegetation and water bodies, representing the natural baseline of the ecosystem.

The Heat Index is derived from Land Surface Temperature (LST) retrieved through remote sensing. Taking Landsat 8 as an example, the thermal infrared band (TIRS) is first used to calculate radiance via the radiative transfer equation, which is then converted to land surface temperature, as shown in Formula (6).

$$LST = \frac{K_2}{\ln\left(\frac{K_1}{L_\lambda} + 1\right)} \quad (6)$$

Among them, K1 and K2 are sensor calibration parameters, L_λ represents radiance, and LST reflects the energy balance and thermal environmental stress of the ecosystem. The following conclusions can be drawn: the LST in urban heat island (UHI) areas and industrial agglomeration zones is significantly higher than in other regions, indicating that human activities continuously intensify surface thermal stress and disrupt the spatial structure of biological habitats and material cycles in ecosystems. In contrast, the LST in natural vegetation areas and water bodies is lower than average, demonstrating their stronger thermal regulation capacity within the ecosystem.

3.3 Calculation and Standardization of RSEI

Due to the significant differences in dimensions and value ranges among NDVI, WET, NDBSI, and LST, normalization is required to compress their values into the [0,1] interval, calculated as Equation (7):

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (7)$$

Among them, X is the original indicator value, and Xmin and Xmax are the minimum and maximum values of the indicator in the study area, respectively. To ensure that all indicators are positively correlated (i.e., the higher the indicator value, the better the ecological quality), it is necessary to normalize and then reverse-process indicators such as NDBSI and LST. The normalized NDVI, WET, NDBSI (after reversal), and LST (after reversal) are incorporated into PCA to extract the first principal component (PC1). PC1 contains the maximum variance information of the original indicators and can comprehensively reflect the main differences in ecological quality. The RSEI is constructed using PC1 and calculated as shown in Equation (8).

$$RSEI = \frac{PC1 - PC1_{min}}{PC1_{max} - PC1_{min}} \quad (8)$$

The final RSEI value ranges between [0,1]. A value closer to 1 indicates better ecological quality, while a value closer to 0 signifies poorer ecological quality. By integrating multiple indicators through PCA, the RSEI achieves a comprehensive quantification of the "greenness-wetness-dryness-heat" interactions within

ecosystems and provides regionalized results for remote sensing quality assessment (Xu and Deng, 2022).

3.4 XGBoost Model

This study employs the XGBoost algorithm to investigate the independent and interactive effects of natural and anthropogenic factors on the dependent variable RSEI. The XGBoost algorithm, proposed by Chen et al. (2016), is an efficient gradient boosting technique (Chen and Guestrin, 2016). As a forward additive model, it sequentially constructs multiple decision trees and combines them into a strong classifier. The final prediction for each sample is obtained by aggregating the results from all individual decision trees (Meng et al., 2016). Using the selected RSEI influencing factors mentioned above, the model is trained to generate an optimal objective function (Obj) for the decision trees, calculated as Equation (9).

$$Obj_j = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma^T \quad (9)$$

In the equation, $G_j = \sum_{i \in I_j} g_i$ represents the sum of the gradients for all samples g_i belonging to the j -th leaf node, where g_i is the first-order derivative of the function; $H_j = \sum_{i \in I_j} h_i + \lambda$, denotes the sum of all samples h_i belonging to the j -th leaf node, where h_i is the second-order derivative of the loss function; T is the number of leaf nodes; λ and γ are hyperparameters. The above formula is used to measure the quality of the tree structure, where a smaller value indicates a better tree structure. The mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) are selected as metrics to evaluate the accuracy of model simulation (Niazkar et al., 2024). The calculation formulas are given by (10), (11), and (12).

$$R^2 = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (12)$$

In the equation, y_i represents the actual value, and y'_i represents the predicted value. The coefficient of determination R^2 ranges from $(-\infty, 1]$, and the closer it is to 1, the better the model's performance, indicating that the predicted results are closer to the true values. MAE and RMSE measure the difference between the model's predicted values and the true values, with smaller values indicating more accurate model results. The training and validation of the model were implemented using Python's XGBoost package.

3.5 SHAP Interpretability Analysis

Although machine learning models excel in predictive performance, their "black box" nature often limits interpretability, making it difficult to uncover the underlying mechanisms between independent and dependent variables. To address this issue and further reveal the intrinsic mechanisms and contribution levels of various natural and human activity factors on RSEI, this study applies SHAP (Shapley Additive Explanations) value analysis. SHAP analysis, proposed by Lundberg et al. (2017), is an interpretable artificial intelligence method based on game theory (Lundberg and Lee, 2017). Its core idea is to provide consistent and reliable local and global explanations for model predictions by calculating the average marginal contribution of each independent variable across all possible feature combinations (i.e., Shapley values).

SHAP values quantify the independent contribution of each variable to the final RSEI prediction for a specific sample (positive values indicate an increase in RSEI, while negative values indicate a decrease). Through aggregated analysis, they can reveal the global importance ranking of variables and their nonlinear relationship patterns with RSEI. Additionally, SHAP interaction values further quantify the strength and direction of the interaction effects between any two independent variables on RSEI, enabling an in-depth exploration of the complex interplay between natural and human activity factors. For detailed SHAP calculation methods, refer to Lundberg and Lee (2017). In this study, SHAP values and their interactions were computed and visualized using the Python shapviz package (Yao et al., 2026).

3.6 PDP Model

The Partial Dependence Plot (PDP) is a model-agnostic global interpretability analysis method that characterizes the marginal effect of one or two features on the model's prediction results, intuitively revealing linear, monotonic, or nonlinear relationships between features and target variables (Friedman, 2001). Its core principle is to calculate the average of the model's prediction results when the target feature takes different values while keeping the distribution of other features unchanged, as mathematically defined in Equation (13).

$$\hat{f}_{x_S}(x_S) = \mathbb{E}_{x_C} [\hat{f}(x_S, x_C)] = \int \hat{f}(x_S, x_C) d\mathbb{P}(x_C) \quad (13)$$

In practical calculations, the Monte Carlo method is usually adopted for approximate estimation, as shown in Formula (14).

$$\hat{f}_{x_S}(x_S) = \frac{1}{n} \sum_{i=1}^n \hat{f}(x_S, x_C^{(i)}) \quad (14)$$

This study combines the results of SHAP feature importance ranking, selects key driving factors to plot PDP curves, analyzes the nonlinear response relationship of each factor to the ecological environmental quality (RSEI) of Yantai City, breaks the "black box" dilemma of the XGBoost model, and provides a more targeted explanatory

basis for the driving mechanism of ecological environmental quality (Friedman, 2001).

IV. ANALYSIS AND RESULTS

4.1 Rationality Analysis of RSEI

The analysis shows that the contribution rate of PC1 ranges from 81.40% to 86.34%, indicating that the first principal component effectively explains the major variations in the ecological environmental quality of Yantai City (Table 2). In terms of load coefficients, NDBSI (-0.68) and NDVI (+0.58) have the largest absolute values, which suggests that the expansion of built-up land and vegetation coverage are the two core factors driving the changes in ecological quality of Yantai. Moreover, this dominant mechanism remained highly stable from 2009 to 2025 (Figure 3).

Additional implications are as follows:

1. The contribution rate of PC1 reaches as high as 80%-86%, proving that RSEI is indeed an efficient ecological evaluation tool.
2. Driving mechanism: NDBSI and NDVI play a dominant role, indicating that the ecological environmental quality of Yantai mainly depends on the interplay between "urbanization construction (negative effect)" and "vegetation restoration (positive effect)".

Table 2 PC1 of RSEI in Yantai City during 2005—2025

Index/Year	2005	2009	2016	2021	2025
NDVI	0.59256	0.5774	0.57534	0.58728	0.59135
WET	0.36562	0.3965	0.39452	0.35905	0.35947
NDBSI	-0.69304	-0.68623	-0.68779	-0.69899	-0.68904
LST	-0.18659	-0.19616	-0.20072	-0.19388	-0.21521
Eigenvalue	0.02327	0.2791	0.03699	0.03939	0.04263
Contribution Rate/(%)	81.4091	84.8255	86.3346	86.1506	85.7708

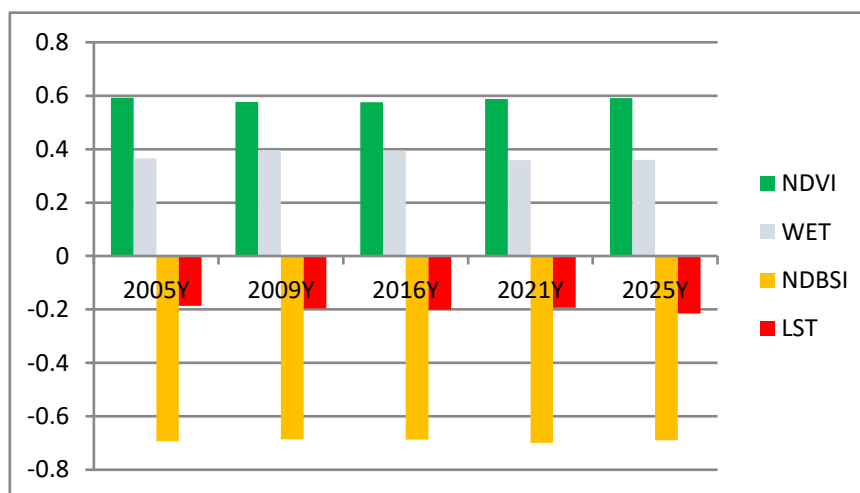


Fig.3 Loading Coefficients of the PC1 of RSEI in Yantai City, 2005—2025

4.2 Spatiotemporal Characteristics of EEQ

4.2.1 Annual Variation of RSEI

In terms of temporal variation, the multi-year average of the RSEI from 2005 to 2025 is 0.438, and the overall EEQ is at the average level (Figure 4), showing a fluctuating trend of initial decline, subsequent rise, another decline and then a further rise. The above periodic fluctuations are closely related to policy orientation in different periods.

From 2005 to 2009, affected by the rapid advancement of urbanization, investment in ecological governance was relatively insufficient, leading to a slight decline in ecological quality. During 2009 to 2016, driven by ecological construction policies such as "Forest City" and

"Water Ecological Civilization City", the ecological quality continued to improve and reached its peak. From 2016 to 2021, with the development of new urban areas and the expansion of industrial land, ecological space was squeezed, resulting in a marked drop in ecological quality. From 2021 to 2025, with the implementation of the "Three Lines and One List" zoning control, ecological restoration projects and green low-carbon development policies, the ecological quality has shown a steady improvement trend. Essentially, the periodic fluctuations of ecological quality in Yantai City stem from the interactive effect of development policies and ecological policies in different stages.

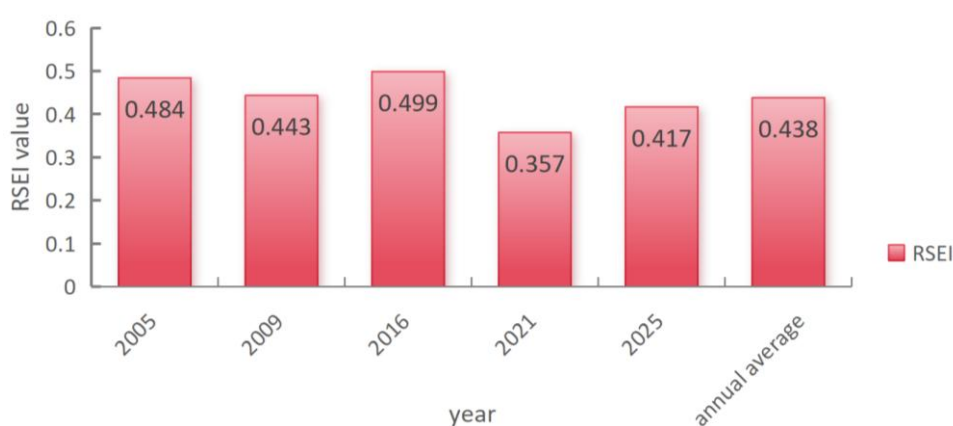


Fig.4 Annual RSEI and Multi-year Mean Value

4.2.2 Spatiotemporal Variation Analysis of RSEI

From 2005 to 2025, the spatial pattern of RSEI in Yantai City exhibited significant phased evolution (Figure

5). In 2005, the overall ecological quality was predominantly moderate, with high-value areas

concentrated in the mountainous and hilly regions of Penglai, Zhaoyuan, Laizhou, and Haiyang. The mountainous areas of Qixia were mainly at a moderate level, with localized low-value patches (Figure 6a). By 2009, ecological quality declined across the entire region, with a significant shrinkage of high-value areas in Qixia, forming contiguous low-value zones in the eastern and western plains and urban areas (Figure 6b). In 2016, ecological quality improved slightly, with the core high-value area in Qixia strengthening and forming a contiguous high-value zone with surrounding regions

(Figure 6c). After 2021, ecological quality declined again, with Qixia becoming the core low-value area of the entire region, forming an east-west contiguous low-value belt with Laizhou and Zhaoyuan (Figure 6d). By 2025, the overall pattern showed a "central collapse and coastal isolation" scenario, with only fragmented high-value patches remaining in the coastal areas of Penglai, Longkou, Laizhou, and Haiyang. Ecological quality returned to a below-moderate level, highlighting the intensifying pressure of human activities on ecological spaces (Figure 6e).

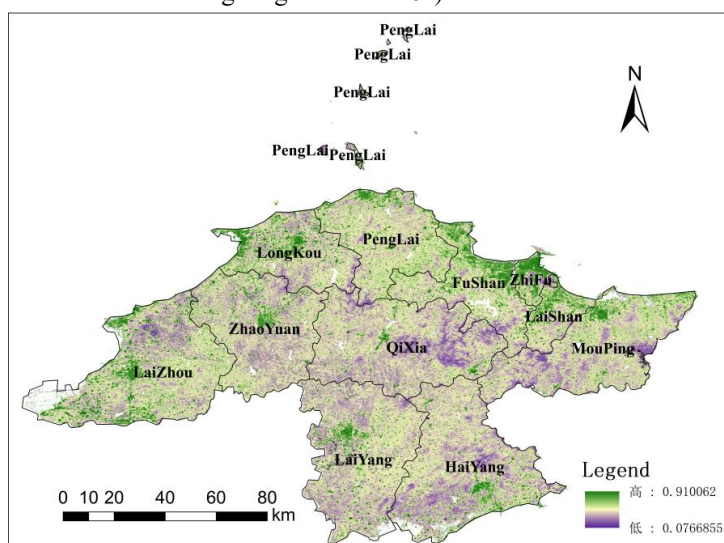
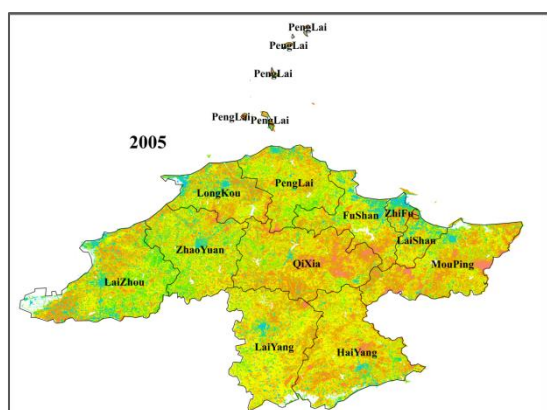
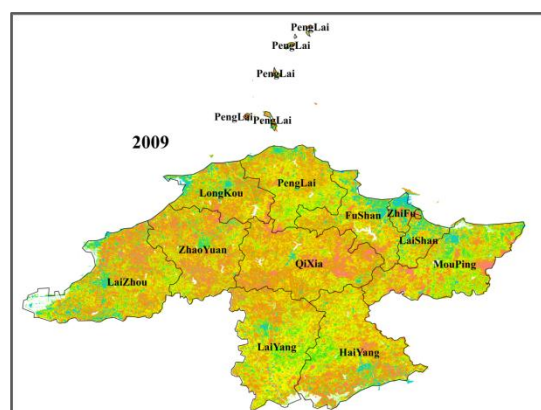


Fig.5 Mean RSEI from 2005 to 2025



(a)2005 RSEI



(b)2009 RSEI

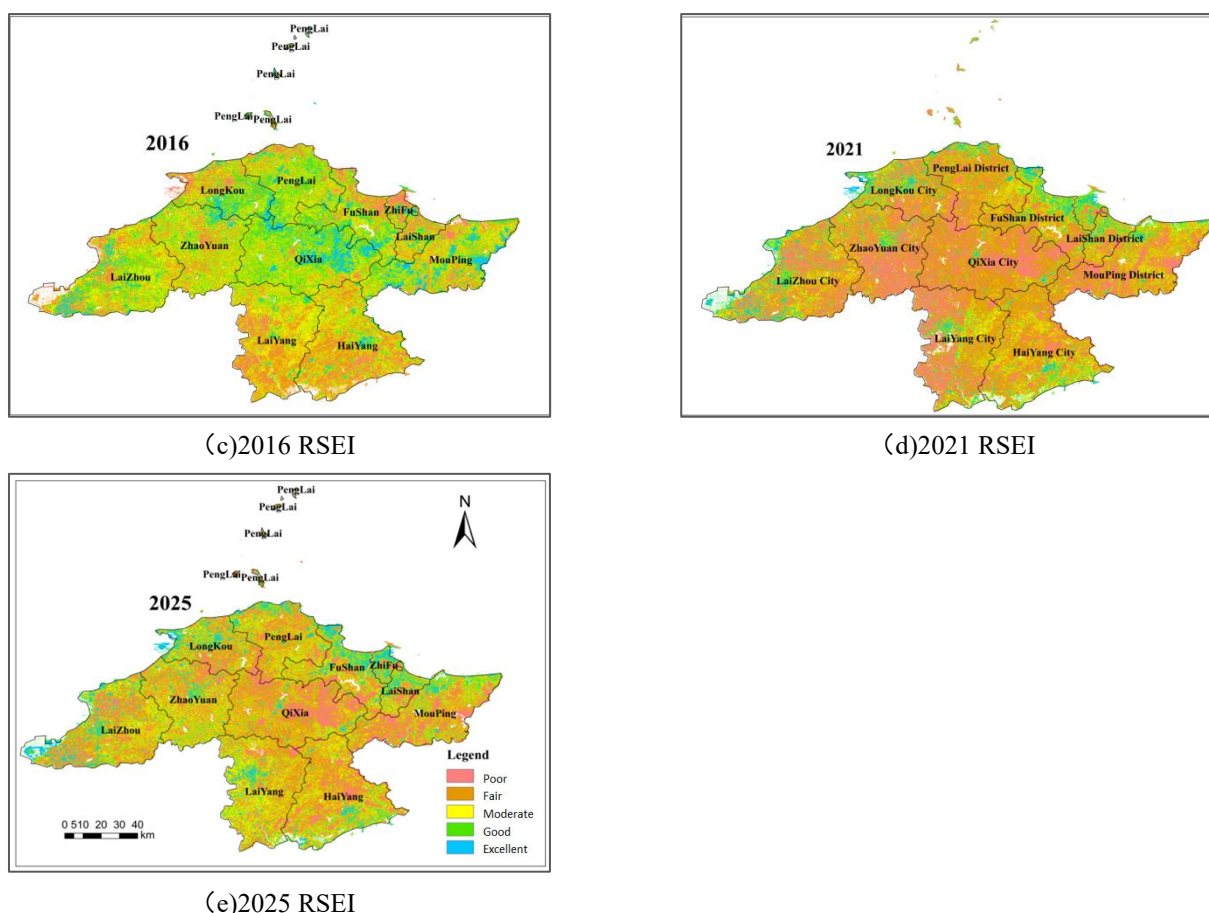


Fig.6 Grading Results of Ecological Environment Quality in Yantai City

From 2005 to 2025, the EEQ of Yantai City underwent a significant phased transformation, characterized by initial deterioration followed by restoration, with an overall positive trend (Figures 7 and 8). In 2005, the ecological quality of Yantai was predominantly at moderate and relatively poor levels, with poor-grade areas accounting for only 3%, relatively poor-grade areas reaching 30%, and moderate-grade areas constituting the highest proportion at 44%. Good and excellent-grade areas accounted for merely 17% and 6%, respectively, indicating that the ecological foundation was primarily moderate, with a relatively low proportion of high-quality ecological space.

With rapid urbanization from 2005 to 2021, the ecological environment faced sustained pressure: both the area and proportion of poor and relatively poor grades increased sharply. By 2021, poor-grade areas rose to 26%, and relatively poor-grade areas reached 42%, together

accounting for nearly 70% of the total. The proportion of moderate-grade areas significantly shrank to 17%, while good-grade areas declined to 10%, and excellent-grade areas remained low. During this period, urban expansion and the encroachment of construction land on ecological spaces, coupled with declining vegetation coverage, led to continuous degradation of ecological quality in central urban areas and northern coastal plains.

From 2021 to 2025, ecological restoration efforts yielded remarkable results, with quality grades showing a comprehensive rebound: the proportion of poor-grade areas decreased from 26% to 17%, and relatively poor-grade areas dropped from 42% to 37%, with both categories experiencing simultaneous reductions in area. The proportion of moderate-grade areas rebounded to 27%, good-grade areas increased to 12%, and excellent-grade areas recovered to 8%, reflecting a continued expansion of high-quality and moderate ecological spaces.

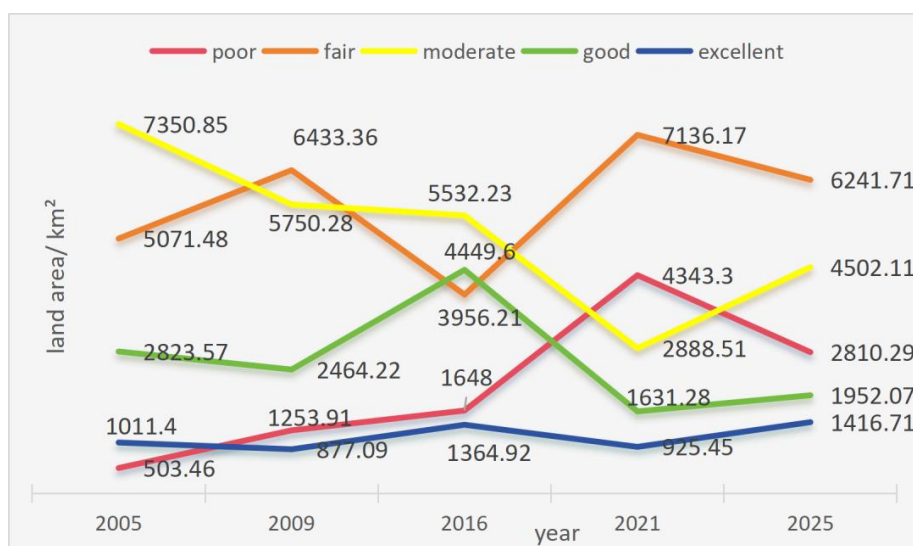


Fig.7 Statistics on Land Area of Each Ecological Environment Quality Grade, 2005–2025

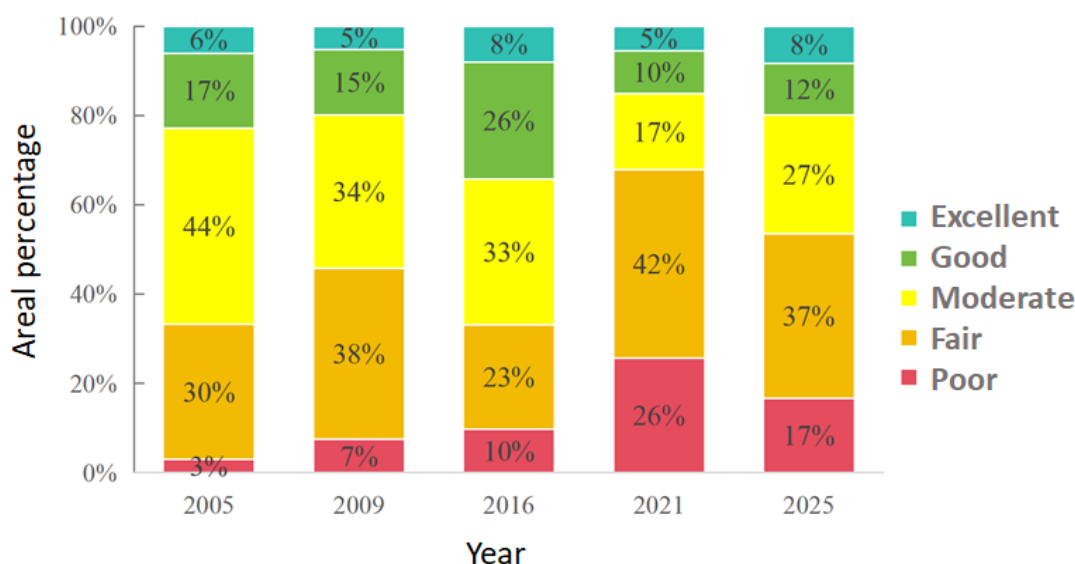


Fig.8 Proportions of Graded Areas of Ecological Environment Quality in Yantai City, 2005–2025

4.3 Analysis of Driving Factors for EEQ

4.3.1 Analysis Based on the XGBoost-SHAP Model

2016 was the peak year for RSEI, exhibiting the most typical spatial differentiation of ecological conditions (Figure 4). Therefore, this year was selected for driving mechanism analysis. The 2016 dataset and influencing indicators were fitted using the XGBoost model in Python, with a random 70%/30% split between the training and testing sets. The model evaluation results showed an R^2 of 0.6485, RMSE of 0.1513, and MAE of 0.151. This R^2 value indicates that the selected factors explain approximately 65% of the RSEI variation, while the

remaining portion may be influenced by factors such as soil type, human policies, local micro-topography, etc.

Figure 9 displays the ranking of the importance of various environmental features in predicting RSEI using the XGBoost model. The results show that built-up land (BL) contributes the most to RSEI prediction, followed by precipitation (PRE), forest land (FL), annual average temperature (AT), and grassland (GL), indicating a strong nonlinear relationship between urbanization intensity and ecological quality in the model. Additionally, elevation (DEM), slope (SLO), population (POP), GDP, and

nighttime light data (NTL) also exhibit significant importance, exerting notable effects on RSEI distribution.

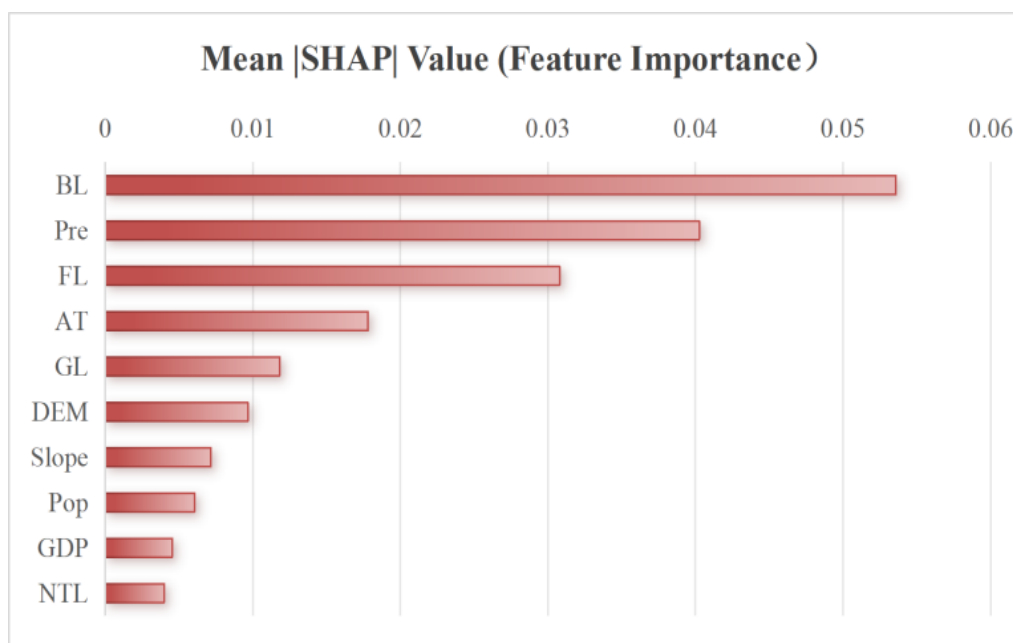


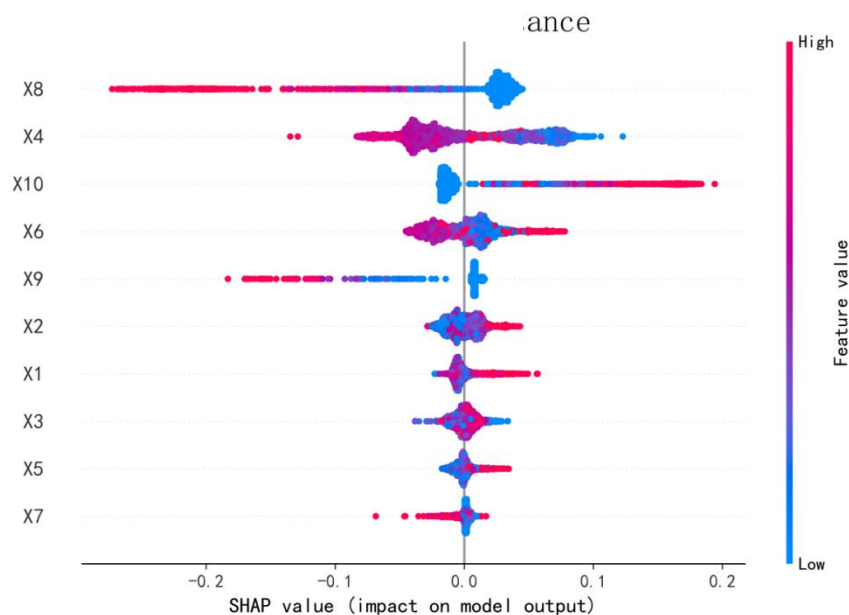
Fig.9 SHAP-based Feature Importance Ranking

Figure 10 further reveals the nonlinear response characteristics of various environmental variables to RSEI predictions. Regions with higher annual precipitation generally correspond to higher RSEI predictions, indicating that sufficient moisture is conducive to improving regional ecological quality. Areas with higher forest coverage also promote RSEI enhancement, which aligns with the regulatory role of water in vegetation growth and surface ecological processes. Construction land has a significant negative impact on RSEI, meaning that regions with a higher proportion of construction land tend to exhibit a declining trend in RSEI, while grassland also shows a negative inhibitory effect on RSEI.

Areas with a higher proportion of construction land tend to have lower RSEI levels, which may be related to ecological degradation processes such as the expansion of impervious surfaces, reduced vegetation cover, and

deterioration of surface thermal environments. Indicators reflecting high-quality vegetation conditions, such as forest cover, generally exhibit a positive influence on RSEI in the SHAP scatter plot, suggesting that vegetation growth status is one of the key controlling factors for the spatial distribution of RSEI. The influence patterns of annual average temperature and elevation on RSEI are more complex, indicating that their relationship with RSEI may be jointly regulated by the interaction of multiple climatic and topographic factors.

The interpretative analysis based on SHAP reveals that land use type, precipitation conditions, and human activity intensity are the dominant environmental drivers influencing the spatial distribution of RSEI in the study area, providing a scientific basis for formulating regional ecological protection and land-use optimization strategies.



Notes: X1: SLO X2: DEM X3: POP X4: PRE X5: GDP X6: AT X7: NTL X8: BL X9: GL X10: FL

Fig.10 SHAP Feature Dependence Distributions

4.3.2 Interpretability Analysis Based on the PDP Model

According to the analysis results of the PDP (Figure 11), the ecological environment quality (RSEI) of Yantai City exhibits significant nonlinear response characteristics to various driving factors. Combined with the ranking results of feature importance, core driving factors with high contribution rates, including built-up land (BL), annual average precipitation (PRE), forest land (FL), and annual average temperature (AT), are selected to further analyze their nonlinear response patterns.

The PDP curve of annual average precipitation (X4) indicates that within the range of 440–480 mm, the RSEI remains at a high level above 0.56 and stays relatively stable. When the precipitation exceeds 480 mm, the RSEI drop rapidly, reaches a trough around 540 mm, and then rebounds slightly but remains significantly lower than the initial level. The results reveal that moderate precipitation can maintain the stability of regional ecosystems, while excessive precipitation is prone to triggering ecological problems such as soil erosion and runoff scouring, thereby leading to the deterioration of ecological environment quality.

The PDP curve of annual average temperature (X6) shows that in the range of 12.6–12.8°C, the predicted RSEI value rises rapidly and peaks near 12.8°C. Within

12.8–13.4°C, the RSEI decreases continuously with the rise of temperature and falls to the minimum value around 13.4°C. After the temperature exceeds 13.4°C, the RSEI presents a rising trend again. This characteristic suggests that appropriate temperature conditions can optimize the regional hydrothermal matching status and promote vegetation growth and the virtuous cycle of ecosystems. Excessively high temperature will intensify evaporative dissipation, induce drought stress, and exert a negative impact on ecological environment quality. In the high-temperature range, as thermal conditions further increase, ecosystems gradually adapt to the high-temperature environment, resulting in a certain rebound of the RSEI.

The PDP curve of built-up land (X8) demonstrates that when the proportion of built-up land is lower than 0.2, the predicted RSEI value stays at a relatively high level above 0.49. With the continuous increase in the proportion of built-up land, the RSEI shows a sustained downward trend; the decline slows down and remains at a low level within the range of 0.6–0.8, and the RSEI plummets sharply when the proportion approaches 1.0. This feature indicates that the expansion of built-up land has a significant negative stress effect on ecological environment quality, and the higher the intensity of urban

development, the more prominent the risk of regional ecological quality degradation.

The PDP curve of forest land (X10) shows that the RSEI is at a low level when the proportion of forest land approaches 0. As the proportion of forest land increases, the RSEI rises rapidly and continuously, with the most obvious growth in the range of 0~0.2, followed by a steady increase and maintenance at a high level in the high-value interval. This characteristic proves that forest land coverage has a significant positive promoting effect on ecological environment quality, serving as a core positive driving factor for the improvement of regional ecological quality. The higher the proportion of forest land, the stronger the stability of the ecosystem.

The short vertical lines at the bottom of Figure 11 mark the distribution density of sample data. The response curves of all factors are mainly concentrated in the sample-dense interval, which verifies the reliability of the analysis conclusions. Overall, PDP analysis reveals the nonlinear response mechanism of ecological environment quality in Yantai City to key driving factors including built-up land, precipitation, forest land and temperature. Among them, low proportion of built-up land, moderate precipitation, high forest land coverage and suitable temperature jointly form the optimal environmental interval for maintaining and improving regional ecological quality.

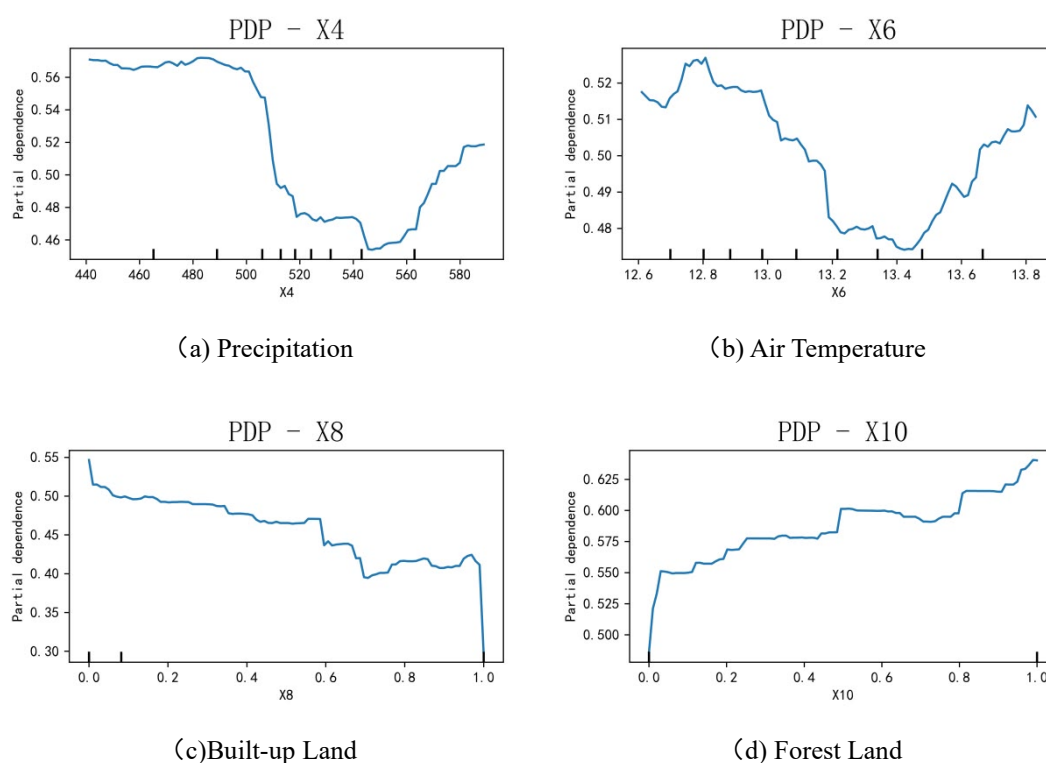


Fig.11 PDP Analysis of Key Environmental Factors

V. CONCLUSION

This study constructs the RSEI of Yantai City from 2005 to 2025 based on GEE platform data, adopts the XGBoost-SHAP model to analyze its influencing factors, and systematically explores the spatiotemporal characteristics of ecological quality in Yantai City during 2005–2025. The main conclusions are as follows:

(1) From 2005 to 2025, the temporal variation of EQQ showed a fluctuating upward trend with an overall moderate grade, experiencing phased changes of "deterioration—improvement—degradation—restoration", and the multi-year average value was 0.438. The phased fluctuation of ecological quality in Yantai City resulted

from the interactive effect of development policies and ecological policies in different periods.

(2) The ecological quality of Yantai City underwent phased changes from 2005 to 2025, eventually forming a spatial pattern of "central depression and coastal isolated high-value zones", with the overall ecological quality declining and the interference effect of human activities intensifying. In terms of spatial differentiation, high-value areas were periodically concentrated in the Qixia mountainous area in the central part, while low-value areas were contiguously distributed in plains and urban built-up areas.

(3) RSEI analysis indicates that NDBSI and NDVI are the core characteristic indicators affecting the ecological quality of Yantai City, corresponding respectively to the negative stress of construction land expansion and the positive promotion effect of vegetation coverage. The XGBoost model further reveals that land use type and annual precipitation are the key factors driving the spatiotemporal evolution of ecological quality. Construction land and forestland serve as the anthropogenic and natural foundations for ecological quality changes by directly altering the surface cover pattern; annual precipitation indirectly regulates the contribution of NDVI to ecological quality by affecting vegetation growth status. The three factors jointly form a complete driving chain of "climate-vegetation cover-anthropogenic disturbance".

Based on GEE and the XGBoost-SHAP model, this study realizes long-time series ecological quality monitoring and interpretable attribution analysis, which can provide scientific support for ecological protection and restoration as well as territorial spatial optimization of coastal cities. The limitations of this study lie in the failure to incorporate variables such as soil type and policy intervention, and the model coefficient of determination R^2 is 0.65. Subsequent research can introduce more feature variables or adopt deep learning models to further improve the model interpretability and explanatory power.

REFERENCES

- [1] Bao, S. W., Yu, C. Z., Wan, Z. Y. Spatial mismatch and attribution analysis of flood risk and resilience in megacities: Insights from a "Risk-Resilience-Effect" framework and an interpretable XGBoost-SHAP model. *Sustainable Cities and Society*, 2025, 131. DOI:10.1016/j.scs.2025.106758.
- [2] Chen, T., and Guestrin, C. XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 785-794.
- [3] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5): 1189-1232.
- [4] Geng, J. Y., Yu, K., Xie, Z., et al. Analysis of spatiotemporal variation and drivers of ecological quality in Fuzhou based on RSEI. *Remote Sensing*, 2022, 14(19): 4900.
- [5] Huang, Y. T., Yu, H. J., Xu, Y. P., et al. Evaluation and analysis of ecological environment quality in Linhai City based on GEE and remote sensing ecological index. *Environmental Science & Technology*, 2024, 47(11): 217-228.
- [6] Li, J. S., Tu, W. Q., and Lu, X. Z. Long-term variation analysis of ecological environment quality in Ningbo City based on remote sensing ecological index. *Urban Geotechnical Investigation & Surveying*, 2025(3): 35-39.
- [7] Liu, Q., Qiao, J. J., Li, M. J., et al. Spatiotemporal evolution of ecological environmental quality and its dynamic relationships with landscape pattern in the Zhengzhou Metropolitan Area: A perspective based on nonlinear effects and spatiotemporal heterogeneity. *Journal of Cleaner Production*, 2024, 480. DOI:10.1016/j.jclepro.2024.144102.
- [8] Lundberg, S. M., and Lee, S. I. A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017: 4768-4777.
- [9] Meng, X. R., Bradley, J., Yavuz, B., et al. MLlib: Machine learning in apache spark. *The Journal of Machine Learning Research*, 2016, 17(1): 1235-1241.
- [10] Niazkar, M., Menapace, A., Brentan, B., et al. Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023). *Environmental Modelling & Software*, 2024, 174. DOI:10.1016/j.envsoft.2024.105971.
- [11] Sun, Q., Qiao, R. T., Jiao, Q. J., et al. Response of ecological quality to land use/cover change during rapid urbanization of Xiong'an New Area. *Land*, 2024, 13. DOI:10.3390/land13122167
- [12] Wang, J. Y., Yang, L. P., Wang, M., et al. Spatiotemporal

- change monitoring and driving force analysis of ecological environment quality in the Loess Plateau of northern Shaanxi. *Environmental Science*, 2025, 46(7): 4522-4533. DOI:10.13227/j.hjlx.202406111.
- [13] Wang, Y. X., Xu, Y. Y., Yang, J. J., et al. Dynamic monitoring and spatiotemporal pattern evolution of ecological environment quality in Chongqing City based on Landsat. *Acta Ecologica Sinica*, 2023, 43(15): 6278-6292.
- [14] Wei, Y. H., Qian, J. P., Fan, W. W., et al. Dynamic monitoring of ecological environment quality in Lijiang River Basin based on RSEI. *Science of Soil and Water Conservation*, 2021, 19(1): 122-131.
- [15] Xu, H. Q. Development and application of remote sensing ecological index in urban areas. *Acta Ecologica Sinica*, 2013, 33(24): 7853-7862
- [16] Xu, H. Q., and Deng, W. H. Rationality analysis of MRSEI index and its difference from RSEI index. *Remote Sensing Technology and Application*, 2022, 37(1): 1-7.
- [17] Xu, Y. H., Zhang, T. B., Yi, G. H., et al. Spatiotemporal characteristics and influencing factors of ecological environment quality in the Taihang Mountain Area. *Environmental Science*, 2025, 46(4): 2428-2438.
- [18] Yang, L. Y., Chen, W. X., Zeng, J., et al. Gradient differences of urbanization impact on habitat quality in the Yangtze River Basin. *Acta Ecologica Sinica*, 2024, 44(10): 4038-4050. DOI:10.20103/j.stxb.202301280140.
- [19] Yao, X., Chen, X., Wang, X. W. Spatiotemporal evolution and driving factors of ecological environment quality in Minnan Delta urban agglomeration based on XGBoost-SHAP model. *Environmental Science*, 2026. DOI:10.13227/j.hjlx.202508029.
- [20] Yuan, B. X., Wu, N., Cheng, P. Spatiotemporal variation characteristics and driving forces of ecosystem types in Chaohu Basin from 2010 to 2020. *Journal of Anhui Agricultural Sciences*, 2022, 50(7): 59-64+77.
- [21] Zhang, Q., Cao, G. C., Zhang, L. L., et al. Dynamic evaluation atlas and driving force analysis of ecological environment quality on the southern slope of Qilian Mountain based on GEE. *Environmental Science*, 2025, 46(7): 4534-4544. DOI:10.13227/j.hjlx.202406195.
- [22] Zhang, X. M., Fan, H. B., Sun, L., et al. Identifying regional eco-environment quality and its influencing factors: A case study of an ecological civilization pilot zone in China. *Journal of Cleaner Production*, 2024, 435. DOI:10.1016/j.jclepro.2023.140308.
- [23] Zhang, Y. F., Yang, Z. H., Zhou, Z. Y., et al. Dynamic assessment and prediction of ecological environment quality in Shantou City based on RSEI and XGBoost model. *China Environmental Science*, 2026. DOI:10.19674/j.cnki.issn1000-6923.20260104.001.
- [24] Zhou, Z., Hu, X. S., Liu, C. Y., et al. Dynamic monitoring and evaluation of ecological environment quality in the upper Yellow River based on GEE: A case study from Longyangxia to Jishixia. *Chinese Journal of Ecology*, 2023, 42(10): 2545-2554. DOI:10.13292/j.1000-4890.202310.034.